Article
ECORFAN Journal
COMPUTING
August 2013 Vol.4 No.10 1061-1066

Study of linear behavior algorithms with fundamental tone detection

SEGURAL-Gabriel. †*´, ALVAREZ-Jesús ´, HERRERA-Juan´´,

´ Instituto Politécnico Nacional, Centro de Innovación y Desarrollo Tecnológico en Cómputo
´´Departamento de procesamiento paralelo, Unidad Profesional Adolfo López Mateos, Av. Juan de Dios Bátiz s/n casi esq. Miguel Othón de Mendizábal edificio del CIDETEC. Colonia Nueva Industrial Vallejo, Delegación Gustavo A. Madero C.P. 07700, México D.F.

Received November 23, 2012; Accepted April 12, 2013

Fundamental tone is the lowest frequency among a set of periodic waves. Determination and calculation will reveal the distance of a nearby object and the object type. Fundamental tone is not an intrinsic characteristic of the object, as not all people and animals make sounds with this frequency, partly within language sound, and not all vowels and letters of the alphabet are fundamental tone. This paper shows a comparison of different techniques for obtaining the fundamental tone of linear behavior by reference to a compact file compression (MP3).

Auto-correlacion, Archive, Cepstrum, Frequency, Fundamental, Language, Wavelet, Waves.

Citation: Segural-G. , Alvarez-J. , Herrera-J. Study of linear behavior algorithms with fundamental tone detection. ECORFAN Journal 2013, 4-10:1061-1066

*Correspondence to Author (email:gsegurab1200@alumno.ipn.mx)
† Researcher contributing first author.

August 2013 Vol.4 No.10 1061-1066

1 Introduction

Fundamental tone or fundamental frequency F0 is determined by the speed of opening or closure of the vocal cords in the larynx during speech sound types [1]. Estimating the frequency has been a significant treatment need from the start of digital signals (voice). Traditionally, fundamental frequency detection was used for the purpose of intonation, currently used in the recognition of tonal languages, conversion of voice, sharp tone recognition fundamental to recognize the different parameters of some pathological conditions. Thus, this is of great interest to calculate the fundamental tone. The detection of fundamental tone is one of the most difficult tasks in digital signal processing because there are audio quality problems, caused by noise level as more serious pronunciation problems are due to medical history in individuals.At present there are different algorithms for the detection of fundamental tone which are classified as:

- Algorithms with understanding loss.
- Compression algorithms without loss.

The lossless compression algorithms allow for some degradation of signal quality to achieve higher levels of understanding as the following formats: MP2, MP3, WMA and others. The lossless compression algorithms for retrieving the signal to its original shape within the formats worked are FLAC, Shorten WAV and others.This research addresses the MP3 file on the extraction of fundamental tone accuracy.

The primary objective is to use the MP3 format for more compact audio understanding to transmit the voice signal in a practical system to organize people with speech pathology.

Section 2 presents the algorithms generally used for the detection of fundamental tone.

II. Fundamnetal algorithms pitch detection.

The algorithms studied are: (Section 2.1) Auto-correlation (Section 2.2) Auto-correlation optimized (2.3) Product harmonic (2.4) ceprstrum and (2.5) Wavelet Transform. Each works with a sampling frequency of 16KHz.In the process of fundamental frequency extracting, the appearance of estimation errors that result in a detected value that is double (Doubling pitch) or half (pitch halving) the appropriate frequency is common [2].

2.1 Auto-correlation.

Auto-correlation is defined as the convolution of the same data within itself, shown in formula 1:

$$R(\tau) = \frac{1}{l}\sum_{j=1}^{L} S_j S_{j-r}$$

```
Auto-correlation algorithm
N = length (Data)
X = data-sample (Data)
Rxx = zeros (1, N)
For m = 1 to N +1
    For n = 1 to N-m +1
Rxx (m) = Rxx (m) + x (n) * x (n + m-1)
    end for
end for
Rxx = Rxx / Rxx (1)
Lag = (0: N-1)
```

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682

The implementation of this algorithm to a digital signal requires at least three periods to ensure a fragment within the signal, which is of 40 ms corresponding to 512 samples for a sampling frequency of 16KHz considering a minimum frequency of 60Hz. Thereby the signal obtained is shown in Figure 2.1.

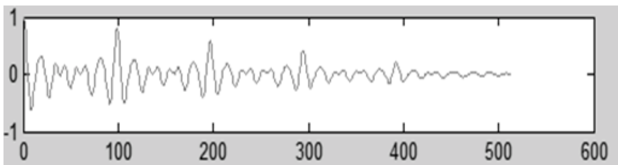


Figure 2.1

The auto-correlation is determined by the distance between the main peak (Energy) and the first secondary maximum, thus calculating the value of the fundamental period T0.

2.2 Optimized Auto-Correlation

The auto-correlation algorithm optimizing auto-correlation is similar to, but with the difference, that it makes a refinement of the estimated values at the fundamental frequency considering non-integer values [3].The Yohav Medan refinement suggests a rational value in the fundamental frequency value found in the auto-correlation algorithm which is shown in formula 2.

$$\gamma_{\gamma}(t,s)=\frac{\sigma^{2}}{2}\left(t^{2\pi}+s^{2\pi}-(s-t)^{2\pi}\right)$$

Autocorrelation algorithm optimized.

N = length (Data)
X = data-sample (Data)
Rxx = zeros (1, N)
For m = 1 to N +1
    For n = 1 to N-m +1
Rxx (m) = Rxx (m) + x (n) * x (n + m-1)
    end for
end for

Rxx = (Rxx) 2 / Rxx (1)
Lag = (0: N-1)

In this way, the selected graph signals a half and a full harmonic avoiding harmonic misinterpretation as shown in Figure 2.2.1.

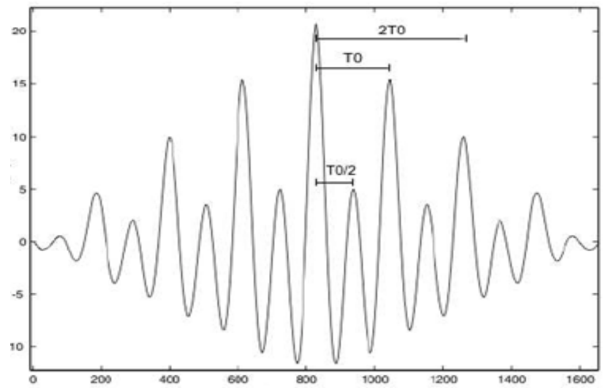


Figure 2.2.1

2.3 Harmonic Product.

The use of spectral techniques to determine, on the frequencies axis, the maximum value corresponding to the fundamental frequency is determined by other fundamental frequencies, as in the case of voice analyzing recommending a minimum frequency of 60Hz. In this section, we have to apply the FFT where the minimum frequency resolution is for low values of the fundamental frequency. In the worst case, when the signal noise bringing the frequency tends to jump, it is advisable to give a range that is 60 Hz to 350 Hz to prevent frequency loss. To resolve a problem, a filling zeros is recommended before applying the FFT [4] [5].
To resolve a problem, a filling zeros is recommended before applying the FFT [4] [5].

$$f_{o}=\frac{M_{i}.F_{S}}{N_{muestras}}$$

Harmonic product algorithm.
N = length (Data)
X = data-sample (Data)
Rxx = zeros (1, N)
For m = 1 to N +1
    For n = 1 to N-m +1
Rxx (m) = Rxx (m) + x (n) * x (n + m-1)

FFT (Rxx (m))
    end for
end for
f_0 = FFT (Rxx (m)) / N

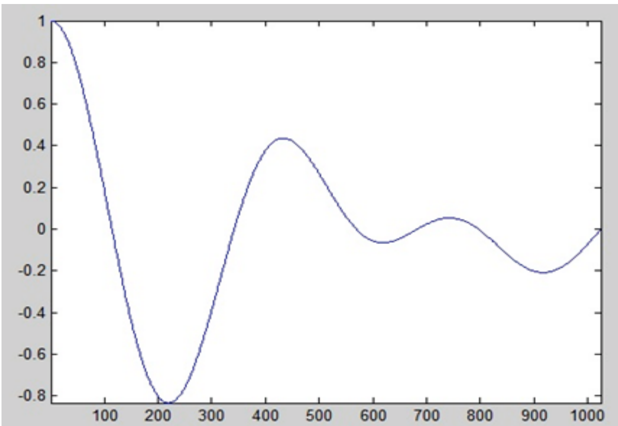


Figure 2.3.1

2.4 Cepstrum

Method for extracting the fundamental frequency, based on the use of the Cepstrum transform. To carry out this process assumes that the audio signal f (t) is the result of impulse response convolution of the vocal tract convolution h (t) with the signal emitted by the glottis s (t).

$$f(t) = h(t) * s(t)$$

The objective of this method is signal de-convolution f (t) to obtain s (t). This is achieved with formula 5:

$$F(w) = H(w)S(w)$$

To perform this procedure first proceeds to decompose the real part imaginary part as shown below.

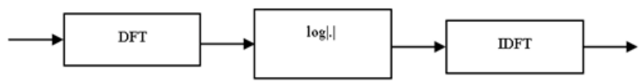
$$\begin{aligned} &FFT(\log|F(w)|) = FFT(log|H(w).S(w)|) = \\ &= FFT(log|\mathrm{H(w)}|+\log|S(w)|) = \\ &= FFT(\log|H(w)|) + FFT(\log|S(w)|) \end{aligned}$$

There fore,the Cepstrum formula 6 can be determined.

$$C = FFT(\log|F(w)|)$$

To determine the fundamental frequency formula 7 is applied.

$$f_0 = \frac{F_s}{q-1}$$



To summarize, a flowchart illustrating the algorithm to obtain a cepstrum is shown in Figure 2.4.1

Figure 2.4.1

X = data-sample (Data)
Dft = real (FFT (Data))
Db = (10. * Log (Data))
Standardization = max (Db);
Fs / 2 = Db / Standardization
db2 = (20/log (10)). * log (data)
cepstrum = real (ifft (db))
max = 0
distance = 0
for N +1 to i
  pitch = max (i)
  distance = max (i) -0
end for
tone = pitch
tone_distance = Distance

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682

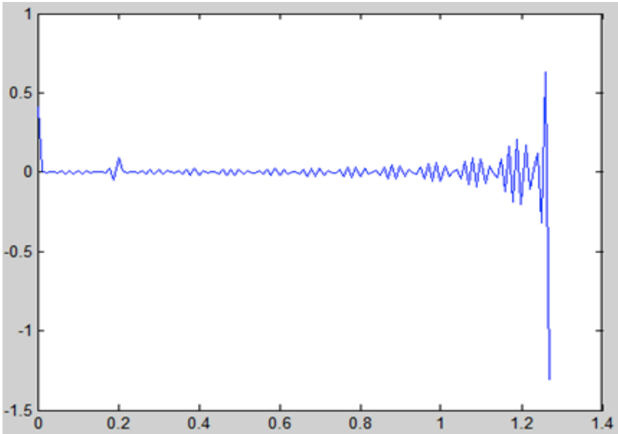


Figure 2.4.2

2.5 Transformed Wavelet.

The first maximum locals selected determine the fundamental frequency allowing the location of the fundamental tones. The phases for this procedure are necessary to perform the following steps [6] [7].

Identifying maximum local: the first stage of the system consists in determining the maximum value for each of the windows; Local maximum correction is performed in two stages; Two maximums located at distance n; The third highest maximum is found which has an average of the two, and hence determines the vector search; Maximum distance. A filter of size N is Applied to eliminate all erroneous frequencies, and is decomposed to determine the distances; Scales selection. From the J vector distances, those p vectors distances that have the least deviation from the mean value are taken.
5. Averaging. The rest is obtained by averaging the fundamental frequency.Applying this procedure obtains a graph similar to that shown in Figure 2.5.1

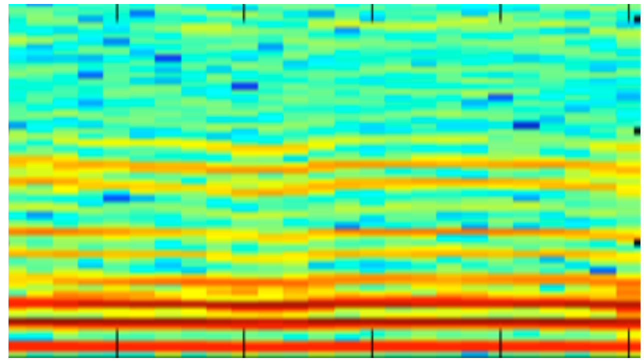


Figure 2.5.1

3. Method

To make the practical implementation of the methods described above, the following steps were developed: a) Initially, the coding and implementation of codes in a simulation environment Matlab was used and subsequently applied to a 16KHz audio sample; b) Methods are used to prove which is best to determine the fundamental tone, showing that the mayoralty show deficiencies when locating the fundamental tone. However, it has better behavior in the wavelet transformer, allowing minor error in locating the fundamental tone y c) It was also observed that not all the methods behave equally when detecting the fundamental tone. There are methods that work best at high frequencies as in the case of auto-correlation and others that work at low frequencies, such as the wavelet that also works at low frequency. In the case of cepstrum it is not known how efficient it is to obtain the fundamental tone as in detecting a noise it cannot detect which is the fundamental tone.

4. Results.

Based on the aforementioned procedure it was observed that, in general terms in the simulation: A good response was observed in the implemented models, in frequency terms, especially in conditions controlled by background noise.

In this part some methods suffer temporary imbalance in the analysis. All this by analyzing a single phoneme 'a' with a value of 16 kHz in a segment of 227 Hz (selected segment) to obtain a table with the following features.

Phoneme	High Frequency	Low Frequency
'a'	730-1090 Hz/ 5 db	2440 Hz /28 db
'A'	640-1190 Hz/ 10 db	2390 Hz /27 db

Table 1

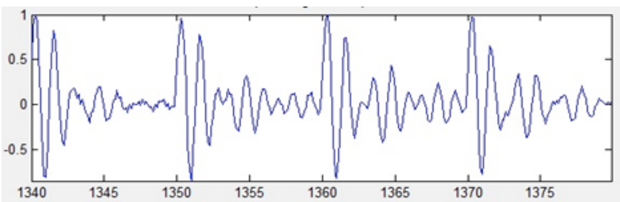


Figure 3.1

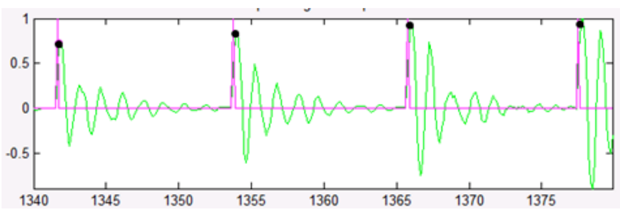


Figure 3.2

5. Conclusions

Based on these results, it was concluded that:

Starting from the structure used to obtain the fundamental tone [8], it was found that the methods implemented, based on frequency domain techniques.
Have superior performances to track the signal in extreme conditions for noise background.
In the same way the wavelet transformer method is employed it has an acceptable behavior and this permits knowing, which is the acceptable fundamental tone.
The wavelet method and algorithm used in digital signal processing allows an average value of the fundamental tones within a full signal.

6.References

C. G. Donald, Speech Processing and Synthesis Toolboxes p 320-324. John Wiley & Sons, Inc. New York USA. 2000.

Good, IJ,The Relationship Between Two Fast Fourier Transforms, Computers, IEEE Transactions on, Vol C-20, (310-317), March 1971.

Hui Ding, A Method Combining LPC- Based Cepstrum and Harmonic Product Spectrum for Pitch Detection, Nanjing University of S&T, China; JiaXing University, China, conference Publications. (537-540) Dec 2006.

Ibon Saratxaga, Iker Luengo, Eva Navas, Inmaculada Hernaez, Jon Sanchez, Iñaki Sainz, Pitch Detection in Adverse Conditions, Aholab - School of Engineering. From the Basque Country University. March 2009.

Medhurst, R.G G. Fundamental and harmonic distortion of waves, frequency- modulated with a single tone, Proceedings of the IEE - Part B: Electronic and Communication Engineering, Vol 107 (155- 164). March 1960.

Rabiner, L. A comparative performance study of several pitch detection algorithms, Bell Laboratories, Murray Hill, NJ, Vol 24, (399-418).

Saito, S. Specmurt Analysis of Polyphonic Music Signals. Audio, Speech, and Language Processing, IEEE Transactions on, Vol 16 (639-650). March 2008.

Zhengmao Ye, Quantitative effects of discrete wavelet transforms and wavelet packets on aerial digital image denoising, Electrical Engineering, Computing Science and Automatic Control,CCE,2009 6th International Conference on, (1-5) Conference Publications, Jan. 2009.